A Section 508–conformant HTML version of this article
is available at https://doi.org/10.1289/EHP7342.

Quantile g-Computation: A New Method for Analyzing Mixtures of Environmental Exposures

Silke Schmidt

https://doi.org/10.1289/EHP7342

Environmental exposures occur in mixtures of similar or different agents from the same or different sources. Attention to mixtures has grown in recent years,[1] as has interest in causal inference techniques.[2,3,4] Statistical methods for mixture analysis, including weighted quantile sum (WQS) regression, have become widely available and used.[5,6] Investigators recently reported in *Environmental Health Perspectives* a new method that combines aspects of WQS with a causal inference method known as g-computation[7] to estimate the joint effect of all exposures in a mixture.[8] The researchers, led by Alexander Keil, call their new method quantile g-computation.

It is difficult to tease apart the potential effects of individual constituents in a mixture. Keil, an assistant professor of epidemiology at the University of North Carolina at Chapel Hill, explains that WQS addresses this problem by reimagining the mixture as a single index. It estimates the effect of an intervention that causes all elements of the mixture to decrease or increase at once, he says. "The 'interventions' idea is important," he explains, "because, as with air pollutants, if we set out to lower one pollutant, it usually results in systemic changes that lead to decreases in multiple pollutants from similar sources."

Once investigators have decided which exposures to include in a mixture of interest, they can ask several possible research questions. WQS regression is designed to answer one of them: It estimates the combined effect of these exposures under the two assumptions that all exposure–outcome associations are *a*) either null or in the same direction, and *b*) linear and additive.

Quantile g-computation relaxes both of these assumptions. In fact, the researchers showed that WQS is a special case of quantile g-computation in large samples if the two assumptions are met. To evaluate the performance of both methods, Keil's team simulated a large data set (sample size of 100,000) and two smaller data sets typical of observational studies (sample sizes of 100 and 500). In some scenarios, WQS produced biased effect estimates when its assumptions were violated.

Of particular interest was the performance of both methods when some mixture components had counteracting effects on an outcome. For the mixture effect estimate, quantile g-computation produced accurate estimates of the true effect, whereas WQS estimates were biased. The new method also produced unbiased estimates when the effects of mixture components were nonlinear. For example, manganese is an important nutrient at low concentrations but can be toxic at high concentrations.[9] Analysts can model this relationship with quantile g-computation but not with WQS.

The researchers also studied the impact of simulated unmeasured confounders on the performance of both methods. Those results

Environmental exposures rarely occur in isolation. But the complexity of chemical mixtures makes it difficult to estimate the effects of individual constituents. Image: © iStockphoto/Alex Potemkin.

were of particular interest to Marianthi-Anna Kioumourtzoglou, an assistant professor of environmental health sciences at Columbia University. When unmeasured confounders were left out, both methods produced biased results. But while the bias increased with a growing number of confounders for WQS, it stayed the same for quantile g-computation.

"This shows that causal inference methods like g-computation don't protect us from bias due to unmeasured confounders," says Kioumourtzoglou, who was not involved in the study. "However, they are more robust and perform better under a wider range of possible models."

Kioumourtzoglou points out that the estimation of nonlinear and nonadditive associations is a desirable feature but requires prior knowledge. For instance, in the manganese example,[9] researchers have to decide *a priori* what type of nonlinear model is appropriate for this and, ideally, any other nutrients in the mixture of interest.

Thomas Webster, a professor of environmental health at Boston University who also was not involved in the study, applauds the authors for promoting a causal framework. "I appreciate that quantile g-computation has a strong theoretical foundation in causality," says Webster. "The study demonstrates that we need to apply mixture methods with caution but also shows that they can be superior to traditional approaches."

Silke Schmidt, PhD, writes about science, health, and the environment from Madison, Wisconsin.

References

1. Taylor KW, Joubert BR, Braun JM, Dilworth C, Gennings C, Hauser R, et al. 2016. Statistical approaches for assessing health effects of environmental chemical mixtures in epidemiology: lessons from an innovative workshop. Environ Health Perspect 124(12):A227–A229, PMID: 27905274, https://doi.org/10.1289/EHP547.
2. Vandenbroucke JP, Broadbent A, Pearce N. 2016. Causality and causal inference in epidemiology: the need for a pluralistic approach. Int J Epidemiol 45(6):1776–1786, PMID: 26800751, https://doi.org/10.1093/ije/dyv341.
3. Weisskopf MG, Kioumourtzoglou MA, Roberts AL. 2015. Air pollution and autism spectrum disorders: causal or confounded? Curr Environ Health Rep 2(4):430–439, PMID: 26399256, https://doi.org/10.1007/s40572-015-0073-9.
4. Howard GJ, Webster TF. 2013. Contrasting theories of interaction in epidemiology and toxicology. Environ Health Perspect 121(1):1–6, PMID: 23014866, https://doi.org/10.1289/ehp.1205889.
5. Christensen KLY, Carrico CK, Sanyal AJ, Gennings C. 2013. Multiple classes of environmental chemicals are associated with liver disease: NHANES 2003–2004. Int J Hyg Environ Health 216(6):703–709, PMID: 23491026, https://doi.org/10.1016/j.ijheh.2013.01.005.
6. Bobb JF, Valeri L, Claus Henn B, Christiani DC, Wright RO, Mazumdar M, et al. 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics 16(3):493–508, PMID: 25532525, https://doi.org/10.1093/biostatistics/kxu058.
7. Snowden JM, Rose S, Mortimer KM. 2011. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. Am J Epidemiol 173(7):731–738, PMID: 21415029, https://doi.org/10.1093/aje/kwq472.
8. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 128(4):47004, PMID: 32255670, https://doi.org/10.1289/EHP5838.
9. Chen P, Bornhorst J, Aschner M. 2018. Manganese metabolism in humans. Front Biosci (Landmark Ed) 23:1655–1679, PMID: 29293455, https://doi.org/10.2741/4665.